

Neural Network based classification of Bio Optical signals for detection of cell abnormalities of human skin

G.Subramanya Nayak
Department of E &C Engineering,
Manipal Institute of Technology
Manipal University
Manipal -576104

Puttamadappa C
Department of E &C Engineering
SJB Institute of Technology
Uttarahalli Road, Kengeri
Bangalore -60

ABSTRACT

Early diagnosis of precancerous and malignant lesions is critical for the improving of the current poor survival rate of patients with a variety of tumors. The development of new high-specificity and high-sensitivity imaging technologies can play an important role in the early diagnosis, accurate staging, and treatment of cancer. Bio-optical signals are the result of the optical functions of the biological systems, occurring naturally or induced by the measurement. The identification of the state of human skin tissues is discussed here. The Bio-optical signals recorded in vitro have been analyzed by extracting various statistical features. Using MATLAB programs, various statistical features are extracted from both normal and pathology spectra. Different features like mean, energy etc were extracted. The values of the feature vector reveal information regarding tissue state. Then a classical multilayer feed forward neural network with back propagation algorithm is employed to serve as a classifier of the feature vector, giving 100% successful results for the specific data set considered. Furthermore, a small time needed to acquire and analyze the fluorescence spectra together with high rates of success, proves our method very attractive for real time applications.

Keywords

Neural Networks, Back Propagation algorithm

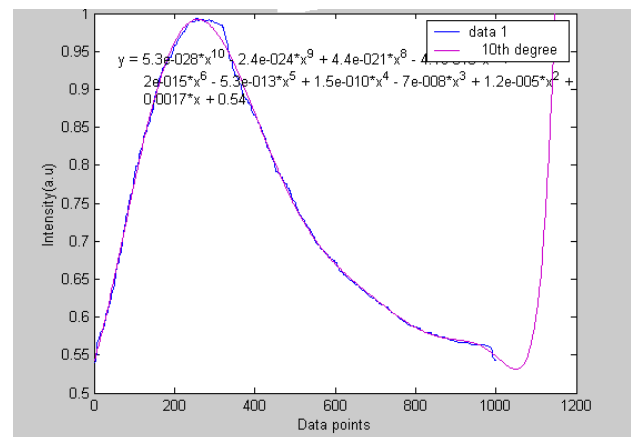
1. INTRODUCTION

The cancer is one of the leading causes of death all over the world, but, if detected early, can be curable. In the current study, the data analysis and classification of pathological conditions of the optical spectra of skin cancer are performed using MATLAB programs. The proposed study is done on MITI-FOPTO database. A large amount of data needs to be analyzed for the classification of normal and pathology conditions. Using MATLAB functions, features like differentiation and count of spikes for different thresholds, mean, standard deviation, energy, residuals on curve fitting, LPC coefficients were extracted. These are used for the classification of spectra in normal and pathological states, using Artificial Neural Network.

Before feature extraction, the signals are normalized and then filtered to eliminate undesirable spikes due to noise and other disturbances.

The 10th degree polynomial curve is fitted onto pathology spectra as shown in Figure 1.

Figure 1. Curve Fitting onto Optical Spectra



This paper deals with the data analysis and classification of pathological conditions of the optical spectra of skin cancer performed using MATLAB.

2. ANN DIAGNOSTIC ARCHITECTURE

An ANN structure is employed for the classification of spectra. It consists of four modules as shown in Figure 2.

The signals are filtered with 21 ordered median filters to eliminate undesirable spikes due to noise and other disturbances. The signal is then fed to the feature extraction module and different features were extracted.

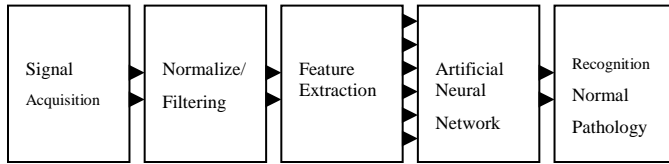


Figure 2. The Model scheme used to perform diagnosis

3. FEATURE EXTRACTION

1) Differentiation

DIFF (X) for a vector X is $[X(2)-X(1) \ X(3)-X(2) \ \dots \ X(n)-X(n-1)]$. DIFF (X) for a matrix X is the matrix of row differences, $[X(2:n, :) - X(1:n-1, :)]$ [13, 14]. 1st order differentiation was carried out on each normal and pathology spectra.

After completion of finding the number of peaks at thresholds 100, 250, 500, 750 for 25 normal and 25 pathology spectra, the mean, standard deviation and variance were determined and are tabulated in Table 3.1 and Table 3.2.

Table 3.1: Number of peaks in normal spectra

Threshold	Mean	Std.deviation	Variance
TH>750	4.045	3.966	9.6193
TH>500	14.5	7.8606	97.2143
TH>250	112.1823	44.8367	1.937×10^3
TH>100	487.0455	62.8906	4.7×10^3

Table 3.1: Number of peaks in arrhythmia spectra

Threshold	Mean	Std.deviation	Variance
TH>750	4	2.7134	2.5714
TH>500	21.7273	4.8617	23.6364
TH>250	178.3636	20.037	401.4805
TH>100	551.8182	24.189	585.1082

2) Energy

For vectors, SUM (X) is the sum of the elements of X, for matrices, SUM (X) is a row vector with the sum over each column that corresponds to energy.

The energy was obtained for different windows of data points for normal and pathology spectra where significant variations in energy distributions were observed.

3) 4th order linear prediction coefficients (LPC)

4th order linear prediction coefficients for normal and pathology spectra are determined and are found as one of the prominent discriminating factor between these two cases. The coefficients like radius and angle are the features that are obtained as per the equation given below.

The radius and the angle in radians are taken from the polar diagram.

4. ANN IMPLEMENTATION

A. Design of ANN with Back Propagation Algorithm:

Three layer networks are sufficient to design any nonlinear network. Input is a layer of nodes with three feature vectors as input. Single second and third layers have same activation function for all neurons. Activation function is a tan sigmoid function as shown in Figure 3.

LM Back propagation algorithm, is used for training.

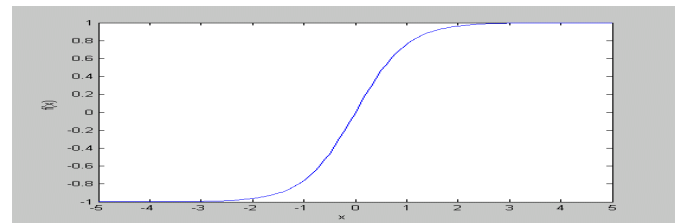


Figure 3. Activation function – Tan sigmoid

B. Training of ANN:

Data from normal and pathology were obtained from data base. Data of 40 normal and 40 pathology are used to train the network.

Feature vector is constructed and training program is executed. The training of ANN stops once the performance goal is met, as shown in Figure 4. The convergence is achieved in 262 epochs.

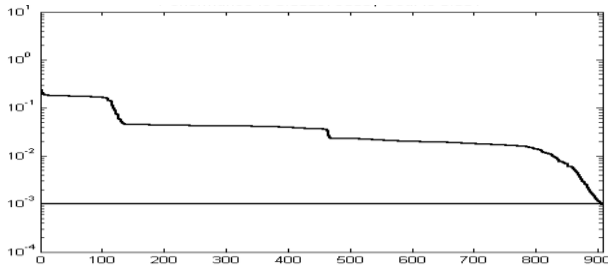


Figure 4. Training of ANN

5. RESULTS

Feature vector of 25 test data of *Normal* and 25 test data of *pathology* are given to network to classify. The results are shown in Table 5.1.

Table 5.1 Classification results

Test data	Desired output	Classifier output	Output Characterization
N1	1	1.0000	Normal
N2	1	1.0000	Normal
N3	1	1.0000	Normal
N4	1	1.0000	Normal
N5	1	1.0000	Normal
N6	1	1.0000	Normal
N7	1	1.0000	Normal
N8	1	1.0000	Normal
N9	1	1.0000	Normal
N10	1	1.0000	Normal
N11	1	1.0000	Normal
N12 – N25	1	1.0000	Normal
P1	-1	-0.9923	Pathology
P2	-1	-0.9008	Pathology
P3	-1	-0.8025	Pathology
P4	-1	-0.9078	Pathology
P5	-1	-0.8070	Pathology
P6	-1	-0.1000	Pathology
P7	-1	-0.8970	Pathology
P8	-1	-0.8033	Pathology
P9	-1	0.9776	Pathology
P10 – P15	-1	-0.9880	Pathology
P16	-1	1.0000	Normal
P17 - P24	-1	-0.9008	Pathology
P25	-1	1.0000	Normal

6. CONCLUSION

Our method achieves the discrimination between normal and pathology optical spectra. Furthermore, a small time needed to acquire and analyze the spectra together with the high rates of success, proves our method very attractive for real-time applications. The results of ANN has been verified with the results of PCA and found satisfactory. This gives us to firmly decide on the pathology. Further, performance parameters can be accurately judged if the tests are carried out on large number of spectra. As less number of features is used, computational delay in training the ANN is reduced. Gradient descent algorithm is too slower and convergence depends on learning rate. Hence high performance algorithm (Levenberg – Marquardt algorithm) which operates in batch mode is used for fast training. Specificity, Sensitivity and Accuracy obtained are 100%, 92 % and 96 % respectively.

REFERENCES

- [1] Jacek M. Zurada , Introduction to Artificial Neural Systems, pp 25–89. Publisher-Pearson Education, 3rd Edition.
- [2] B.K.Manjunath, J.Kurien, C.Muralikrishna, Autofluorescence of Oral tissue for optical pathology in Oral malignancy, Journal of Photochemistry and Photobiology B. Biology 73 (2004) 49 -58.
- [3] George A.Rovithakis and Michail Maniadakis ,Artificial Neural Networks for discriminating Pathologic from Normal Peripheral Vascular Tissue, IEEE transactions on Biomedical Engineering ,vol 48, No 10, october 2001
- [4] Sigurdur Sigurdsson et ,al ,Detection of skin cancer by classification of Raman spectra, IEEE transactions on Biomedical Engineering ,2004.
- [5] Simon Hykin, Neural Networks-A Comprehensive foundation, section 4.1to4.10, page 156-234, Jaico Publishing House, 5TH edition.
- [6] William Palm J, “Introduction to MATLAB 6 for Engineers”.
- [7] G.S.Nayak, Sudha Kamath et al. (2006) Principal Component Analysis and Artificial Neural Network Analysis of oral Tissue Fluorescence Spectra: Classification of Normal, Premalignant and Malignant Pathological Condition Biopolymers 82:152-166
- [8] G.Subramanya Nayak and Gopallakrishna Prabhu K (2004) Analysis of Electroglottographic Signal for Measurement of Vocal Fold Abduction / Larynx Movement ,using Amplitude Modulation Technique, ECCS-2004 Proc. National Conference on Electronic Circuits and Communication Systems, Patiala, India, 2004, pp 463 – 46